DOCUMENT RESUME

ED 228 316 TM 830 242

AUTHOR Blumberg, Carol Joyce; And Others
TITLE Comparison of Methods of Data Analysis in Nonrandomized Experiments.
PUB DATE Apr 83
NOTE 31p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, Montreal, Quebec, April 11-15, 1983). Research supported in part by a grant from the University of Delaware Research Foundation.
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; *Control Groups; *Data Analysis; Data Collection; Evaluation Methods; Mathematical Models; *Research Design; Research Methodology; *Simulation
IDENTIFIERS *Monte Carlo Studies; *Nonrandom Selection

ABSTRACT

Various methods have been suggested for the analysis of data collected in research settings where random assignment of subjects to groups has not occurred. For the purposes of this paper the set of allowable nonrandomized designs is made up of those research designs where data are collected for one or more groups of subjects at two or more time points on some measure of interest. Further, none of the groups need be a control group. The main purpose of the paper is to describe and report the results of a Monte Carlo simulation study that was carried out to determine which of several data analysis methods developed by either Blumberg and Porter or Olejnik yields the best point estimates of treatment effects under various constraints. When growth on the measure of interest is linear over time, Blumberg and Porter's methods provide the best estimates. When growth is exponential over time, the results are mixed: under some constraints Olejnik's method is best, but usually Blumberg and Porter's methods provide the best estimates. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED228316

Comparison of Methods of Data Analysis in Nonrandomized Experiments

Carol Joyce Blumberg

Sigurd L. Andersen, Jr.

Roberta E. G. Murphy

Linda D. Waters

University of Delaware

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

C J Blumberg

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Printed in U.S.A.

TM 830 242

Paper presented at the 1983 Annual Meeting of the American Educational Research Association, Montreal, April 1983.

The research for this paper was supported in part by a grant from the University of Delaware Research Foundation to Carol Joyce Blumberg.

Abstract

Various methods have been suggested for the analysis of data collected in research settings where random assignment of subjects to groups has not occurred. For the purposes of this paper the set of allowable nonrandomized designs is made up of those research designs where data are collected for one or more groups of subjects at two or more time points on some measure of interest. Further, none of the groups need be a control group. The main purpose of the paper is to describe and report the results of a Monte Carlo simulation study that was carried out to determine which of several data analysis methods developed by either Blumberg and Porter or Olejnik yields the best point estimates of treatment effects under various constraints. When growth on the measure of interest is linear over time Blumberg and Porter's methods provide the best estimates. When growth is exponential over time the results are mixed: under some constraints Olejnik's method is best, but usually Blumberg and Porter's methods provide the best estimates.

Various methods have been suggested for the analysis of data collected in research settings where random assignment of subjects to groups has not occurred. For the purposes of this paper the set of allowable nonrandomized designs is made up of those research designs where data are collected for one or more groups of subjects at two or more time points on some measure of interest. Further, none of the groups need be a control group. The designs making up this set are most often referred to as either nonequivalent control group designs and/or interrupted time series designs (Cook & Campbell, 1979). The main purpose of this paper is to report the results of a Monte Carlo simulation study that was carried out by the authors to determine which of the several data analysis methods to be described in the next section results in the best point estimators of treatment effects under the various conditions studied.

Data Analysis Methods

All of the data analysis methods to be compared in this paper assume some type of continuous natural growth model which is supposed to describe the changes(i.e., growth) in the measure of interest over time. Blumberg (along with Porter) has developed several methods for deriving point estimates of treatment effects (Blumberg, 1982a; Blumberg, 1982b; Blumberg & Porter, 1982). All of these methods assume the

following model of growth over time:

$$Y^{*}_{ij}(t) = g_j(t) \cdot Y^{*}_{ij}(t_1) + h_j(t) + \alpha_j(t)$$

and (1)

$$Y_{ij}(t) = Y^{*}_{ij}(t) + e_{ij}(t) \ ;$$

where $Y^{*}_{ij}(t)$, $Y_{ij}(t)$ and $e_{ij}(t)$ represent the true scores, observed scores, and errors of measurement, respectively, for the ith individual in the jth group, on the measure of interest;

$g_j(t)$ and $h_j(t)$ are continuous functions;

$\alpha_j(t)$ represents the population treatment effect for the jth group;

and t_1 is an arbitrary time point.

Further assumptions are:

(1) Classical measurement theory holds. That is, for each time t, $Y^{*}_{j}(t)$ and $e_j(t)$ are uncorrelated and $E(e_{ij}(t)) = 0$.

and (2) Treatment effects are additive. The expression $g_j(t) \cdot Y^{*}_{ij}(t) + h_j(t)$ represents the natural growth portion of this class of models and represents any natural growth situation where there is a correlation within each group between true scores at any two points in time. Finally, the treatment effects, as defined by the $\alpha_j(t)$'s in the system of equations (1), are not the same as the usual definition of treatment effects.

Let $\alpha(t)$ be the grand mean of the $\alpha_j(t)$'s. The usual definition of a treatment effect is given by $\alpha_j(t) - \alpha(t)$.

All that is required in order apply Blumberg and Porter's methods is that the functional forms of the $h_j(t)$'s are known (e.g., $h_1(t)$ is a logarithmic function of the form $h_1(t) = \log_b(c\cdot(t-t_1) + 1)$, where b and c are co ,cants, possibly unknown; $h_2(t)$ is a linear function of the form $h_2(t) = c\cdot(t-t_1)$ where c is some constant, possibly unknown; etc). In this paper three of Blumberg and Porter's methods will be described and used in the simulation study. The reason for not discussing the remainder of their methods is that the remaining methods are not applicable under the conditions imposed for this particular simulation study.

Blumberg and Porter's first method requires that the data analyst have knowledge of the functional forms of the $g_j(t)$'s and $h_j(t)$'s. It further requires that pretest observations under natural growth conditions are available on the measure of interest at at least M time points, where M is the maximum of (i) two more than the number of unknown constants in the functional form expression for $g_j(t)$; and (ii) two more than the number of unknown constants in the functional form expression for $h_j(t)$. For convenience, this method will be called Method A and p will denote the number of pretest time points.

If one considers the system of equations (1) for each of the p pretest points, then the structural model depicted in Figure 1 can be set up relating the pretest observations at the various time points. In this figure and in the remainder of the paper, without loss of generality, the j subscript representing group membership is dropped. This structural equations model contains many unknown parameters, namely $g(t_2)$, $g(t_3)$, ... , $g(t_p)$, $h(t_2)$, $h(t_3)$, ... , $h(t_p)$, the variance of the true scores at time t_1, and the variances of the errors of measurement at t_1, t_2, ..., t_{p-1}, and t_p.

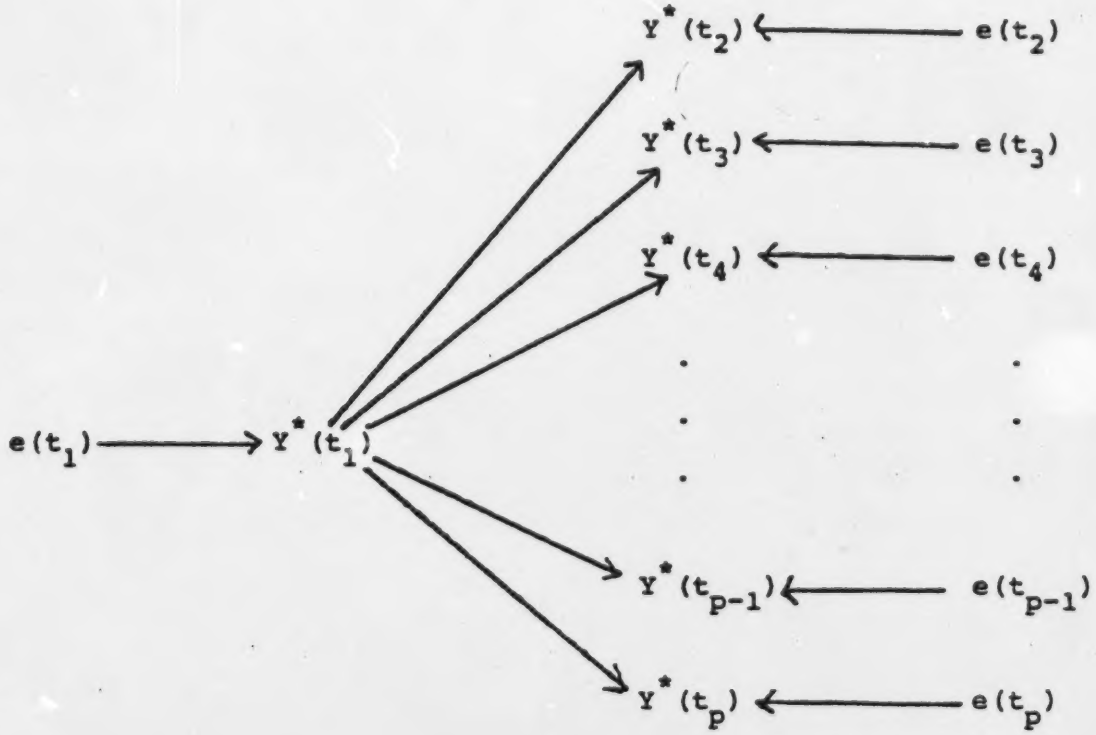

Figure 1

Pictorial representation of the structural model

To implement Method A it is necessary to obtain maximum likelihood estimates of these unknown parameters. But, the structural model is overidentified and hence does not have a closed solution for the maximum likelihood estimates of the parameters. Consequently, LISREL (Jöreskog & Sörbom, 1978) or some other maximum likelihood structural equations computer program must be used to obtain the maximum likelihood estimates. The Appendix gives the LISREL IV input stream corresponding to Figure 1. Let $\widehat{g(t_k)}$ represent the obtained maximum likelihood estimate of $g(t_k)$ for $k = 2,3,\ldots, p$. The maximum likelihood estimates of the $h(t_k)$'s for $k=2,3,\ldots,p$ are obtained by using $\widehat{h(t_k)} = \overline{Y(t_k)} - \widehat{g(t_k)} \cdot \overline{Y(t_1)}$. The $\widehat{g(t_k)}$'s and the $\widehat{h(t_k)}$'s thus provide estimates for the true values of $g(t)$ and $h(t)$, respectively, at the pretest time points. The method of least squares is then used to obtain estimates of the unknown constants in the functional form expressions for $g(t)$ and $h(t)$. For example if $g(t) = b \cdot c^{(t-t_1)} + (1 - b)$, then the estimates of b and c are those values which minimize the quantity

$$\sum_{k=2}^{p} (g(t_k) - (b \cdot c^{(t-t_1)} + (1 - b)))^2 .$$

New functions, labelled $\hat{g}(t)$ and $\hat{h}(t)$ are formed by substituting the estimates of the unknown constants, that were obtained using the process just described, back into the functional form

expressions for g(t) and h(t). For example, if $g(t) = b \cdot c^{(t-t_1)} + (1 - b)$ and $\hat{b} = 1.45$ and $\hat{c} = 2.7$, then $\hat{g}(t) = 1.45 \cdot (2.7)^{(t-t_1)} - .45$. Point estimates of treatment effects are finally given under Method A by

$$\widehat{\alpha_A(t)} = \overline{Y(t)} - (\hat{g}(t) \cdot \overline{Y(t_1)} + \hat{h}(t)) .$$

Blumberg and Porter's second method, to be called Method B, depends upon assuming that the reliability of Y, the measure of interest, is constant over time and upon having knowledge of the exact nature of h(t) [e.g., knowing that $h(t) = 3 \cdot t$ or $h(t) = \log_3[4(t-1)^7]$, etc.] and requires observations at only one pretest time point, namely t_1. Under Method B, point estimates of treatment effects are given by, where $S_Y(t)$ represents the standard deviation of Y(t),

$$\widehat{\alpha_B(t)} = \overline{Y(t)} - \left\{ \frac{S_Y(t)}{S_Y(t_1)} \cdot \overline{Y(t_1)} + h(t) \right\}$$

Blumberg and Porter's third method, to be called Method C, depends upon assuming that the reliability of Y is constant over time and upon having knowledge of the functional forms of g(t) and h(t). Further, both g(t) and h(t) can each only have one unknown constant (e.g., $g(t) = b \cdot (t-t_1) + 1$ and $h(t) = \log_3[c \cdot (t-t_1)^7]$). Method C also requires that pretest observations are available at two pretest time points, say t_1 and t_2. This method is a combination of some aspects of Methods

A and B. Under Method C, $g(t_2)$ can be estimated by $\frac{S_Y(t_2)}{S_Y(t_1)}$. Call this estimator of $g(t_2)$ by the name $\widehat{g(t_2)}$. The value of $h(t_2)$ is then estimated by using $\widehat{h(t_2)} = \overline{Y(t_2)} - \widehat{g(t_2)} \cdot \overline{Y(t_1)}$. The equations $\widehat{g(t_2)} = g(t_2)$ and $\widehat{h(t_2)} = h(t_2)$ are then solved for the unknown constants. These solutions provide estimators of the unknown constants. For example, if $g(t) = b \cdot (t - t_1) + 1$, then the equation $\frac{S_Y(t_2)}{S_Y(t_1)} = b \cdot (t_2 - t_1) + 1$ is solved for b yielding $\hat{b} = \left\{\frac{S_Y(t_2)}{S_Y(t_1)} - 1\right\}/(t_2 - t_1)$. Once the estimates of the unknown constants are obtained, new functions labelled g(t) and h(t) are formed, as in Method A, by substituting the estimates of the unknwon constants into the functional form expressions for g(t) and h(t). Point estimates of treatment effects are then given by

$$\widehat{\alpha_C(t)} = \overline{Y(t)} - [\hat{g}(t) \cdot \overline{Y(t_1)} + \hat{h}(t)] .$$

Olejnik (1977) assumes the following model for the mean population growth over time on the measure of interest:

$$\mu_Y(t) = [b \cdot (t - t_1) + 1] \cdot \mu_Y(t_1) + c \cdot (t - t_1) + \alpha(t)$$

and

$$Y_i(t) = Y_i^*(t) + e_i(t) ,$$

where $\mu_Y(t)$ is the population mean for $Y^*(t)$. Hence, Olejnik requires that the population mean natural growth over time be

linear while Blumberg and Porter allow natural growth to follow any continuous function. Olejnik, however, does not require the assumption of a correlation of +1 between true scores at any two points in time, as is required by Blumberg and Porter's model of natural growth. Olejnik's method, as did Blumberg and Porter's Method C, requires observations to be available at exactly two pretest time points, namely t_1 and t_2. Under Olejnik's method, the point estimators of treatment effects are given by

$$\widehat{\alpha_0(t)} = \overline{Y(t)} - \overline{Y(t_1)} - [\overline{Y(t_2)} - \overline{Y(t_1)}] \cdot \frac{t - t_1}{t_2 - t_1} .$$

All of the four methods just described have some unsolved problems associated with them. The methods developed by Blumberg and Porter are based on maximum likelihood estimation and/or the use of ratios of standard deviations. Both maximum likelihood techniques and estimators based on ratios of standard deviations are known to often lead to biased, although consistent, estimators. One unsolved problem is whether each of Blumberg and Porter's methods lead to estimators whose bias is at an acceptable or unacceptable level. Further, nothing is known about the standard errors of the estimators generated by these methods. Olejnik's method has only been studied when the population natural growth pattern was taken to be linear over time. It can easily be shown by elementary algebra and

statistics that when population mean growth is linear that Olejnik's method produces unbiased estimates of treatment effects. Olejnik (1977) studied the standard error of his method for linear mean population growth under various constraints on the errors of measurement. The bias and standard error of Olejnik's method have not, however, been studied for non-linear mean population growth. The computer simulation study to be described presently, thus, had several purposes:

(i) to study the bias of Blumberg and Porter's Methods A, B, and C and Olejnik's method under various natural growth formulations;

(ii) to study the standard errors of the four methods;

(iii) to compare the estimates obtained under the four methods;

and (iv) to make recommendations for the use of these methods on real data sets.

There is only one other class of methods known to the authors by which one can obtain point estimates of treatment effects. This class of methods which was developed by Strenio, Bryk, and Weisberg (Bryk, Strenio, & Weisberg, 1980; Strenio, Weisberg, & Bryk, in press) is based on the ideas of Empirical Bayes estimation. The use of their class of methods demands a great deal of mathematical and statistical sophistication on the part of the data analyst. Hence, even though Strenio, Bryk, & Weisberg

have produced an excellent class of methods, their methods were not included in this simulation study because of their complexity.

Set Up of Simulation Study

One thousand two hundred data sets were generated in the following manner. First, the canned program NRAN31 was used to generate 13 standard normal random deviates for each of 25 individuals. This program and all remaining programs mentioned in this paper were run on the Burroughs 7700 computer at the University of Delaware. A base true score for each individual was established by adding 5 to the first standard normal random deviate generated for each individual. Without loss of generality, this time point was set equal to $t = 1$. Two different sets of individuals' true scores under natural growth over time were generated at 11 additional time points, which were taken to be equally spaced at t=2,3, ..., 11, and 12 , using $Y^*(t) = g(t) \cdot Y^*(1) + h(t)$ where $g(t)$ and $h(t)$ were certain specified functions. The first set of true scores was generated by setting $g(t) = .5(t - 1) + 1$ and $h(t) = .3 \cdot (t - 1)$. The second set of true scores was generated using $g(t) = .7 \cdot (1.2)^{t-1} + .3$ and $h(t) \equiv 0$. Next, it was assumed that the reliability of Y was constant across time. Three different values were taken for this reliability: .5 , .7 , and .9 . For each of these reliability values the

second through thirteenth standard normal random deviates generated at the first step were used to add on errors of measurement to the true scores in order to generate observed scores with the required reliability values. Thus, for each set of 25 individuals, six different data sets were generated. The properties of the six data sets are enumerated below:

(1) The first data set, which will be referred to as .5 Linear, was generated using $g(t) = .5(t-1) + 1$, $h(t) = .3 \cdot t$, and a reliability of .5 .

(2) The second data set, which will be referred to as .7 Linear, was generated using $g(t) = .5(t-1) + 1$, $h(t) = .3 \cdot t$, and a reliability of .7 .

(3) The third data set, which will be referred to as .9 Linear, was generated using $g(t) = .5(t-1) + 1$, $h(t) = .3 \cdot t$, and a reliability of .9 .

(4) The fourth data set, which will be referred to as .5 Exponential, was generated using $g(t) = .7 \cdot (1.2)^{t-1} + .3$, $h(t) \equiv 0$, and a reliability of .5 .

(5) The fifth data set, which will be referred to as .7 Exponential, was generated using $g(t) = .7 \cdot (1.2)^{t-1} + .3$, $h(t) \equiv 0$, and a reliability of .7 .

(6) The sixth data set, which will be referred to as .9 Exponential, was generated using $g(t) = .7 \cdot (1.2)^{t-1} + .3$, $h(t) \equiv 0$, and a reliability of .9 .

The procedure just described in the preceding paragraph was repeated 200 times yielding a total of 1200 simulated data sets. Notice that when the data sets were generated no treatment effects were entered into the data. Hence, when the four methods described in the last section are used to estimate treatment effects, the calculated values of the estimated treatment effects do in fact represent the bias in the methods because the theoretical values of all treatment effects were set to zero.

For the .5 Linear, .7 Linear, and .9 Linear data sets the estimates of treatment effects for the various methods were calculated in the following manners. For Method A the time points t=1,2,3,4,5, and 6 were taken as the pretest time points. The simulated observed scores for each data set corresponding to these six time points were entered into the LISREL program illustrated in the Appendix. The LISREL estimates of GA(1,1), GA(2,1), GA(3,1), GA(4,1), and GA(5,1) were then used as the maximum likelihood estimates of g(2), g(3), g(4), g(5), and g(6), respectively. It was then assumed that $g(t) = b\cdot(t - 1) + 1$ and that $h(t) = c\cdot(t - 1)$. The method of least squares was then used to estimate b and c. In this case, because both g(t) and h(t) are linear, closed expressions for $\hat{b}$ and $\hat{c}$ are available and are given by $\hat{b} = (\hat{Q} - 15)/55$ and by $\hat{c} = (\overline{Y(2)} + 2\overline{Y(3)} + 3\overline{Y(4)} + 4\overline{Y(5)} + 5\overline{Y(6)} - \hat{Q}\cdot\overline{Y(1)})/55$, where

$\hat{Q} = \widehat{g(2)} + 2\cdot\widehat{g(3)} + 3\cdot\widehat{g(4)} + 4\cdot\widehat{g(5)} + 5\cdot\widehat{g(6)}$. Finally, $\widehat{\alpha_A(t)}$ was calculated for t=7,8,9,10,11, and 12 using $\widehat{\alpha_A(t)} = \overline{Y(t)} - (\hat{b}(t-1) + 1)\cdot\overline{Y(1)} - \hat{c}\cdot(t-1)$. For Method B it was assumed that $h(t) = .3\cdot t$ (the correct function) and t = 1 was taken as the only required pretest time point. To calculate the $\widehat{\alpha_B(t)}$'s the formula $\widehat{\alpha_B(t)} = \overline{Y(t)} - ((S_Y(t)/S_Y(1))\cdot\overline{Y(1)} + .3\cdot t)$ was used for t=2,3,4,5,6, and 7 . For Method C and for Olejnik's method it was assumed that $g(t) = b\cdot(t-1) + 1$ and $h(t) = c\cdot t$ and the pretest time points were taken as being t=1 and t=2. When these linear functions are assumed for g(t) and h(t), Method C and Olejnik's method result in the same estimates for treatment effects. For ease of later discussion, these estimates will be referred to as the estimates from Olejnik's method and are given by $\widehat{\alpha_O(t)} = \overline{Y(t)} - \overline{Y(1)} - (\overline{Y(2)} - \overline{Y(1)})\cdot(t-1)$ for t=3,4,5,6,7, and 8.

For the .5 exponential, .7 exponential, and .9 exponential data sets the estimates of the treatment effects were calculated in the following manners. For Method A the time points t=1,2, 3,4,5, and 6 were taken as the pretest time points and the $\widehat{g(t)}$'s for t=2,3,4,5,and 6 were generated using LISREL as described in the previous paragraph. It was then assumed that $g(t) = b\cdot c^{(t-1)} + (1-b)$. Since h(t) was set to be identically equal to zero when generating the data sets, h(t) was assumed to be identically

equal to zero for Method A and for all the other methods when simulating the analysis methods for the exponential data sets. The method of least squares, using the ZXSSQ subroutine of the IMSL package, was then employed to estimate b and c, and $\widehat{\alpha_A(t)}$ was calculated using $\widehat{\alpha_A(t)} = \overline{Y(t)} - [\hat{b}\cdot\hat{c}^{(t-1)} + (1 - \hat{b})]\cdot\overline{Y(1)}$ for t=7,8,9,10,11,and 12 . For Method B, t = 1 was used as the pretest time point and $\widehat{\alpha_B(t)}$ was calculated using the formula $\widehat{\alpha_B(t)} = \overline{Y(t)} - \frac{S_Y(t)}{S_Y(1)}\cdot\overline{Y(1)}$ for t=2,3,4,5,6, and 7. Method C is not applicable for the exponential data sets since $g(t) = b\cdot c^{(t-1)} + (1 - b)$, which is the corresponding functional form for the g(t) used to generate the data sets, has two unknown constants. For Olejnik's method the time points of the pretests were taken as t=1 and t=2 and the formula $\widehat{\alpha_O(t)} = \overline{Y(t)} - \overline{Y(1)} - (\overline{Y(2)} - \overline{Y(1)})\cdot(t - 1)$ for t=3,4,5,6,7,and 8 was still used to estimate the treatment effects, even though it was realized that Olejnik's assumption of population mean growth being linear does not hold for these data sets.

Results and Conclusions

The easiest way to report the results of this simulation study is by the use of tables. Tables 1 through 6 give the results for the .5 Linear, .7 Linear, .9 Linear, .5 Exponential,

.7 Exponential, and .9 Exponential data sets. As was mentioned

---

Insert Tables 1 to 6 Here

---

earlier, the observed mean for each of the estimators over the 200 simulated data sets is the same as the observed bias of these estimates since the theoretical value of the treatment effects is zero. This observed bias is reported in each table in the column labelled Observed bias. The standard deviation of each of the various estimated treatment effects over the 200 data sets is an estimate of the standard error of the estimators and is reported in the column of each table labelled standard deviation. In each table the column labelled Percentage best represents the number of times that the indicated method yielded an estimated treatment effect whose absolute value was less than the absolute value of the estimated treatment effects generated using the other two methods. Conversely, the column labelled Percentage worst reports the number of times that the indicated method yielded an estimated treatment effect whose absolute value was more than the absolute value of the estimated treatment effects generated using the other two methods. The rows labelled A and B refer to Blumberg and Porter's methods and the rows labelled O refer to Olejnik's method. The starred values in Tables 4,5, and 6 are crude estimates of the observed

bias and standard deviations rather than the actual values. To keep the computer programming tractable, values of estimated treatment effects which were smaller than -1000 were treated as missing when the observed bias and standard deviations were computed. Hence the observed biases are even more negative than indicated and the standard errors are even bigger than indicated. The reason for including the crude estimates of bias and standard deviation is that they do give an indication of the problems associated with Method A when exponential growth is used.

From inspection of Tables 1 through 6 several conclusions can be drawn. When g(t) and h(t) are linear (Tables 1,2, and 3), Method A leads to point estimates of treatment effects which appear to have no noticeable bias while Method B leads to biased estimates. Olejnik's method theoretically leads to unbiased estimates and this was confirmed by the simulation study. Method B has much larger standard errors than either Method A or Olejnik's method. Further, Method B, for all reliability levels and for all posttest time points, rarely gives estimates with smaller absolute value (i.e., Percentage best is lower) than either Method A or Olejnik's method and, in fact, most often, yields the estimates with the largest absolute value (i.e., Percentage worst is high). Hence, Method B can be eliminated as a possible method for analyzing data which follow a linear growth pattern over time. Therefore, the choice of data analysis methods for linear growth is reduced to Method A and Olejnik's method.

Since both Method A and Olejnik's method lead to virtually unbiased estimators, the choice between them must be made based on considerations other than bias. When one extends one time point beyond the last pretest for all three reliability levels the standard error for Olejnik's method is less than the standard error for Method A and further, Olejnik's method leads to smaller absolute values of estimated treatmetn effects a larger percentage of the time. When one extends two time points beyond the pretest for all three reliability level the standard errors and Percentages best and worst are approximately the same for both methods. When one extends three or more time points beyond the pretests for all three reliability levels the standard errors and Percentages worst are smaller and the Percentages best are larger for Method A than for Olejnik's method. Hence, for measures of interest whose true growth pattern over time is linear, it appears that if one wants to extend only one time point beyond the pretests that Olejnik's method should be used. If one wants to extend two time points beyond, it appears to be a toss-up. But, Olejnik's method is much simplier to use and hence is recommended when extending two time points beyond the pretests. When one wants to extend 3 or more time points beyond the pretests, Method A appears to be the preferrable method.

When the true growth pattern on the measure of interest follows the exponential growth model (Tables 4, 5, and 6)

Method A can immediately be eliminated as a possible data analysis method because of its huge standard errors. Hence, when data follow an exponential growth model the choice of data analysis methods is limited to Method B and Olejnik's method. For all 3 reliability levels and when one extends any number of time points beyond the pretest time points Method B appears to give estimates of treatment effects with no noticeable bias while Olejnik's method always leads to biased estimators. The biasedness of Olejnik's method should not, however, be surprising since the method assumes linear growth and the growth model used to generate the data was not linear. When one extends only one or two time points beyond the pretest time points for all three reliability levels Olejnik's method has smaller standard errors and Percentages worst and larger Percentages best than does Method B. Hence, despite being slightly biased, Olejnik's method appears to be the preferrable method when extending only one or two time points beyond the pretests.

When extending three time points beyond the pretest time points the choice of methods is dependent on the reliability of the data. When the reliability is .9, Method B has a larger Percentage best and smaller Percentage worst than Olejnik's method. Further Method B provides virtually unbiased estimates while Olejnik's method leads to biased estimates. Hence when extending three time points beyond the pretests, with data of reliability of .9, Method B is the preferred

method, even though Olejnik's method has a smaller standard error. When the reliability of the data is either .5 or .7 and one is extending three time points beyond the pretests, the Percentages best and worst are almost identical for Method B and for Olejnik's method. As mentioned earlier, Method B leads to virtually unbiased estimates while Olejnik's method leads to biased estimates. However, the standard errors associated with Olejnik's method are smaller. So, when the reliability of the data is either .5 or .7 and one is extending three time points beyond the pretests, data analysts must decide whether they want unbiasedness, in which case Method B should be chosen, or smaller standard errors, in which case Olejnik's method should be chosen. When extending four or more time points beyond the pretests Method B becomes the recommended method for all three reliability levels. The reason for this recommendation is that Method B remains virtually unbiased while the bias inherent is Olejnik's method becomes larger as the data is extended more and more time points beyond the pretests. Further, the Percentages best are larger and the Percentages worst are smaller for Method B than for Olejnik's method.

Limitations and Directions for Further Research

This simulation study was carried out using only two forms of natural growth models--a linear growth model and an exponential

growth model. Further, for each of the two models only one set of parameters was used to generate the data. Also, the assumption of equal reliability across time was made. Hence, the results reported in this paper are very limited. On a positive note, however, the results of this simulation study show that Blumberg and Porter's methods and Olejnik's method are viable data analysis methods. This is important because these methods are not well known and hence have rarely been used in actual data analysis situations. Because of the limitations just cited, much further research needs to be done. First, for the functional forms studied here, parameter values other than those used in this study should be investigated. Second, functional forms other than linear or exponential should be studied. Third, constraints on the errors of measurement other than that of equal reliability should be included in future simulation studies. Finally, this study only concerned the point estimation of treatment effects. Both Blumberg and Porter (Blumberg, 1982a; Blumberg, 1982b; Blumberg & Porter, 1982) and Olejnik (1977) have developed interval estimation and hypothesis testing procedures based on their point estimation procedures. The utility of their interval estimation and hypothesis testing procedures still needs to be studied.

References

Blumberg, C. J. Methods for data analysis in nonequivalent control group designs. American Statistical Association 1982 Proceedings of the Social Statistics Section, pp. 197-202, 1982a.

Blumberg, C. J. The estimation and hypothesis testing of treatment effects in nonequivalent control group designs when continuous growth models are assumed. Unpublished Ph.D. dissertation, Michigan State University, 1982b.

Blumberg, C. J., & Porter, A. C. The estimation and hypothesis testing of treatment effects in nonequivalent control group designs when continuous growth models are assumed. American Educational Research Association annual meeting, March 1982 (ERIC Document ED 220 516/TM 820 533).

Bryk, A. S., Strenio, J. F., & Weisberg, H. I. A method for estimating treatment effects when individuals are growing. Journal of Educational Statistics, 1980, 5, 5-34.

Cook, T. D., & Campbell, D. T. Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally, 1979.

Joreskog, K. G., & Sorbom, D. LISREL IV: Analysis of linear structural relationships by the method of maximum liklihood: User's guide. International Educational Services, 1978.

Olejnik, S. F. Data analysis strategies for quasi-experimental studies where differential group and individual growth rates are assumed. Unpublished Ph.D. dissertation, Michigan State, 1977.

Strenio, J. F., Weisberg, H. I., & Bryk, A. S. Empirical Bayes estimation of individual growth curve parameters and their relationship to covariates. Biometrics, in press.

Appendix

LISREL IV Input Stream

Title Card
DA NG=1 NI=p NO=*sample size* MA=CM
LABELS
*
'YT2' 'YT3' 'YT4' ... 'YTP' 'YT1'
RA
*
Data is next using the following form
$Y_1(t_2)\ Y_1(t_3)\ \ldots\ Y_1(t_p)\ Y_1(t_1)$
$Y_2(t_2)\ Y_2(t_3)\ \ldots\ Y_2(t_p)\ Y_2(t_1)$
.
.
.
$Y_N(t_2)\ Y_N(t_3)\ \ldots\ Y_N(t_p)\ Y_N(t_1)$
MO NY=p-1 NX=1 NE=p-1 NK=1 LY=ID LX=ID BE=ID C
GA=FU,FR PH=DI,FR PS=ZE TE=DI,FR TD=DI,FR
ST *Give starting values for GA(1,1) to GA(p-1, 1),PH(1), TE(1) to TE(p-1), and TD(1) making sure that they are all positive.*
OU MR FD SE ND=8

Table 1

Results for .5 Linear

Number of time points beyond last pretest	Method	Observed bias	Standard deviation	Percentage best	Percentage worst
1	A	.074	0.937	35.5 %	21 %
	B	.296	2.547	11.5	69
	O	.014	0.738	53	10
2	A	.064	1.073	44.5	13
	B	.308	3.490	11.5	71.5
	O	-.005	1.098	44	15.5
3	A	.131	1.219	43	13.5
	B	.620	5.176	14.5	68.5
	O	.108	1.543	42.5	18
4	A	.001	1.375	50.5	9
	B	.546	6.452	10.5	76.5
	O	-.053	2.015	39	14.5
5	A	-.182	1.479	54.5	7.5
	B	1.002	7.972	8	74
	O	.094	2.293	37.5	18.5
6	A	.027	1.677	56.5	7.5
	B	1.034	8.653	6.5	77.5
	O	.087	2.657	37	15

Table 2

Results for .7 Linear

Number of time points beyond last pretest	Method	Observed bias	Standard deviation	Percentage best	Percentage worst
1	A	.049	0.613	37.5 %	15 %
	B	.227	1.983	8.5	76
	O	.009	0.483	54	9
2	A	.042	0.703	45.5	10
	B	.223	2.740	9	77
	O	-.003	0.719	45.5	13
3	A	.086	0.798	48	9.5
	B	.486	4.035	9	75.5
	O	.070	1.010	43	15
4	A	.001	0.900	53.5	7.5
	B	.455	5.122	7.5	81
	O	-.034	1.319	39	11.5
5	A	-.119	0.968	55.5	6
	B	.728	6.180	6	81.5
	O	.062	1.501	38.5	12.5
6	A	.018	1.098	53.5	7.5
	B	.887	6.773	10	80.5
	O	.057	1.739	36.5	12

Table 3

Results for .9 Linear

Number of time points beyond last pretest	Method	Observed bias	Standard deviation	Percentage best	Percentage worst
1	A	.025	0.312	37.5 %	16 %
	B	.122	1.145	6.5	75.5
	O	.005	0.246	56	8.5
2	A	.021	0.358	46	8.5
	B	.102	1.619	8.5	80.5
	O	-.002	0.366	45.5	11
3	A	.044	0.406	49	10
	B	.263	2.338	6	79.5
	O	.036	0.514	45	10.5
4	A	.000	0.458	55	5.5
	B	.252	3.030	4.5	88
	O	-.018	0.672	40.5	6.5
5	A	-.061	0.493	54	4.5
	B	.349	3.560	6	84
	O	.031	0.764	40	11.5
6	A	.009	0.559	56.5	5.5
	B	.532	3.972	6	84
	O	.029	0.886	37.5	10.5

Table 4

Results for .5 Exponential

Number of time points beyond last pretest	Method	Observed bias	Standard deviation	Percentage best	Percentage worst
1	A	-11.562	55.541	3 %	86.5 %
	B	-0.064	1.226	36.5	11
	O	0.153	0.561	60.5	2.5
2	A	-57.052*	388.770*	4	88
	B	-0.170	1.487	38.5	7
	O	0.455	0.855	57.5	5
3	A	-72.352*	429.307*	3	89
	B	-0.098	1.750	46.5	7.5
	O	1.037	1.203	50.5	3.5
4	A	-121.758*	467.494*	4	87
	B	-0.255	1.991	57.5	5
	O	1.700	1.582	38.5	8
5	A	-178.686*	766.859*	4.5	86
	B	-0.107	2.096	71.5	4
	O	2.834	1.845	24	10
6	A	-218.836*	956.396*	5	86
	B	-0.214	3.092	75.5	3.5
	O	4.225	2.162	19.5	10.5

Table 5

Results for .7 Exponential

Number of time points beyond last pretest	Method	Observed bias	Standard deviation	Percentage best	Percentage worst
1	A	-4.993	27.927	3.5 %	85.5 %
	B	-0.045	0.985	28	11.5
	O	0.149	0.367	68.5	3
2	A	-138.164*	963.686*	5.5	85.5
	B	-0.133	1.199	36	12
	O	0.453	0.560	58.5	2.5
3	A	-57.992*	664.887*	4	87.5
	B	-0.062	1.402	48	8
	O	1.009	0.788	48	4.5
4	A	-46.826*	366.526*	4	90.5
	B	-0.166	1.569	61	2.5
	O	1.702	1.037	35	7
5	A	-74.629*	438.340*	5.5	88.5
	B	-0.095	1.697	80	2
	O	2.804	1.212	14.5	9.5
6	A	-83.196*	407.534*	4.5	86.5
	B	-0.093	2.418	85	1
	O	4.194	1.423	10.5	12.5

Table 6

Results for .9 Exponential

Number of time points beyond last pretest	Method	Observed bias	Standard deviation	Percentage best	Percentage worst
1	A	-3.011	21.806	5.5 %	79 %
	B	-0.016	0.598	23	19
	O	0.144	0.187	71.5	2
2	A	-117.414*	811.526*	7	85
	B	-0.070	0.721	42	9
	O	0.983	0.403	51	6
3	A	-3.878*	27.861*	8	83.5
	B	-0.016	0.843	64.5	6
	O	0.983	0.403	27.5	10.5
4	A	-15.978*	123.460*	5.5	85
	B	-0.066	0.939	86	1
	O	1.704	0.532	8.5	14
5	A	-62.224*	560.774*	5.5	85
	B	-0.052	1.029	91.5	0.5
	O	2.776	0.626	3	14.5
6	A	-67.189*	409.017*	5	84.5
	B	0.014	1.405	93.5	0
	O	4.165	0.742	1.5	15.5